**SIGNAL PROCESSING OF NETWORKS（GRAPHS）**

**Xiao Jin**

**Degree Stream : Electronic Engineering**

**Project Supervisor : Dr. Klara Stokes**



An Interim Project Report submitted in partial fulfilment of the requirements of the Maynooth University for the degree of Bachelor of Engineering

**Feb 2020**

**Department of Electronic Engineering**

**Maynooth University**

Table of Contents

[1 Introduction: 3](#_Toc510089126)

[2 Problem Statement and Objectives of the projects: 3](#_Toc510089127)

[3 Background: 3](#_Toc510089128)

[3.1 History of Gait: 3](#_Toc510089129)

[3.2 Modern Gait Analysis: 5](#_Toc510089130)

[3.2.1 Accelerometers, Gyroscopes and Magnetometers: 6](#_Toc510089131)

[3.2.2 Force Sensitive Resistors: 8](#_Toc510089132)

[3.2.3 Flexible Goniometer: 8](#_Toc510089133)

[4 Lit Review: 9](#_Toc510089134)

[4.1 Report on Gyroscopes for Gait Analysis: 10](#_Toc510089135)

[4.2 Report on Xsens MVN NIOMECH: 11](#_Toc510089136)

[4.3 Report on Accelerometer Test-Retest Reliability 12](#_Toc510089137)

[5 Work Plan and Gantt CHART: 13](#_Toc510089138)

[6 References 17](#_Toc510089139)

[Figure 1-Results of Edward Muybridge’s gait experiment. 3](#_Toc475431421)

[Figure 2-block diagram of gait analysis with cameras. 4](#_Toc475431422)

[Figure 3-Accelerometer Block Diagram 5](file:///C:\Users\fyp2017\Google%20Drive\FYP\IntermReport\IntermReport.docx#_Toc475431423)

[Figure 4-Mems of Accelerometer 5](file:///C:\Users\fyp2017\Google%20Drive\FYP\IntermReport\IntermReport.docx#_Toc475431424)

[Figure 5-examples of Flexible goniometers 6](file:///C:\Users\fyp2017\Google%20Drive\FYP\IntermReport\IntermReport.docx#_Toc475431425)

[Figure 2-layout of sensors in the shoe 7](#_Toc475431426)

[Figure 3-Joint angle waveforms and CMC values during one representative level walking trial. 9](#_Toc475431427)

[Figure 6-Gantt chart part 1 Overview of project layout. 13](#_Toc475431428)

[Figure 7-Gantt Chart part 2 30/January - 20/March 14](#_Toc475431429)

[Figure 8-Gantt Chart part 3 17/March - 6/May 14](#_Toc475431430)

[Figure 9-4/May - 24/May 14](#_Toc475431431)

**ABSTRACT**

Deep learning is the most important part of machine learning. In many related fields, the method of deep learning deep neural network is the most advanced solution. In order to further understand the reason why deep neural network is superior in solving practical problems.

This project proposes to use the method of graph signal processing to detect the characteristics of the intermediate representation of deep neural network. By comparing different characteristics and measurement methods, it shows that k-nearest neighbor graph can be used to explain the reason why deep neural network has good performance.

**Declaration**

*I declare that this report and the project it describes is my original work only. I have not plagiarized or excessively quoted the work of others, nor have I colluded with others to represent collaborative work as my own. I confirm that I have appropriately cited all information derived from the published and unpublished work of others.*

*Signed: Student Number: Date:*

# Introduction:

As a branch of machine learning, deep neural network has been paid more and more attention by researchers. Many problems that cannot be solved in traditional theory can be solved well by using deep neural network. The basic structure of the deep neural network is to establish a multi-layer parameter structure between the input and output nodes to connect. Obviously, when there are enough layers of the deep neural network, the corresponding mathematical function obtained can greatly approximate the natural relationship between the input signal and the output signal. There are two opinions on why deep neural network is so magical. One is that deep neural network is a universal approximator of continuous function in vector space. The other is that although deep neural network needs a lot of parameters, it can be effectively trained by error back propagation mechanism. [1]

Although the parameters of the deep neural network can be converged gradually through back propagation, there are still some obvious defects in the deep neural network at present:

1. If the amount of data is not large enough, the accuracy of neural network is relatively poor.[2]
2. Even if the amount of data in the training set is large enough, if the quality of the training set is not high enough, the generalization ability of the neural network will be poor, that is, over fitting will occur.[3][4]
3. At present, there is no clear method to effectively find the hyperparameter in the deep neural network.

The reason why deep neural network has these problems is that deep neural network is like a ‘black box’, so it is necessary to give an effective theory to explain the function of deep neural network. In fact, the deep neural network is composed of directed acyclic graph, and graph signal processing can extend the Fourier analysis to any topological domain described by graph. Therefore, this paper proposes the method of using graph signal processing (GSP) to monitor the middle layer of the training process of the deep neural network, detect the fitting of the deep neural network through GSP, and try to find the deep neural network The feature changes in the process of network performance reaching the best, which provides guidance for the construction and training of deep neural network.

# Problem Statement and Objectives of the projects:

The problem statement for this project is: “How to design a suitable Deep Neural Networks and choose several representative features such that can monitor the DNN well”. The project has been broken down into the following main objectives:

1. Design a Deep Neural Network to do some tasks such as classification.
2. Training the above DNN by the typical data set such as CIFAR-10 to make the accuracy achieve an stable value.
3. Define several representative features about the process of the DNN training.
4. Use GSP method to observe the change of the above features.

These objectives provide a linear progression to the project and help to keep the project on track. These objectives will be discussed in further detail in the work plan section of the report.

# Background:

In this section the various aspects of the background to the project will be presented.

## History of GSP:

Data is all around us, and massive amounts of it. Almost every aspect of human life is now being recorded at all levels: from the marking and recording of processing inside the cells starting with the advent of fluorescent markers, to our personal data through health monitoring devices and apps, financial and banking data, our social networks, mobility and traffic patterns, marketing preferences, and many more. The complexity of such networks [1] and interactions means that the data now reside on irregular and complex structures that do not lend themselves to standard tools.

Graphs offer the ability to model such data and complex interactions among them. For example, users on Twitter can be modeled as nodes while their friend connections can be modeled as edges.

This paper explores adding attributes to such nodes and modeling those as signals on a graph; for example, year of graduation in a social network, temperature in a given city on a given day in a weather network, etc. Doing so requires us to extend classical signal processing concepts and tools such as Fourier transform, filtering, and frequency response to data residing on graphs. It also leads us to tackle complex tasks such as sampling in a principled way. The field that gathers all these questions under a common umbrella is graph signal processing (GSP) [2], [3]. While the precise definition of a graph signal will be given later in the paper, let us assume that a graph signal is a set of values residing on a set of nodes. These nodes are connected via (possibly weighted) edges. As in classical signal processing, such signals can stem from a variety of domains; unlike in classical signal processing, however, the underlying graphs can tell a fair amount about those signals through their structure. Different types of graphs model different types of networks that these nodes represent. Typical graphs that are used to represent common real world data include Erdo˝s–Rényi graphs, ring graphs, random geometric graphs, small-world graphs, power-law graphs, nearest-neighbor graphs, scale-free graphs, and many others. As in classical signal processing, graph signals can have properties, such as smoothness, that need to be appropriately defined. They can also be represented via basic atoms and can have a spectral representation. In particular, the graph Fourier transform allows us to develop the intuition gathered in the classical setting and extend it to graphs; we can talk about the notions of frequency and bandwidth limitation, for example. We can filter graph signals. They can be sampled, with GSP, We can denoise graph signals, we can learn their underlying structure, we can model them. If the graphs cannot be directly observed, we can also learn their structure from data.

## History of Deep Neural Network:

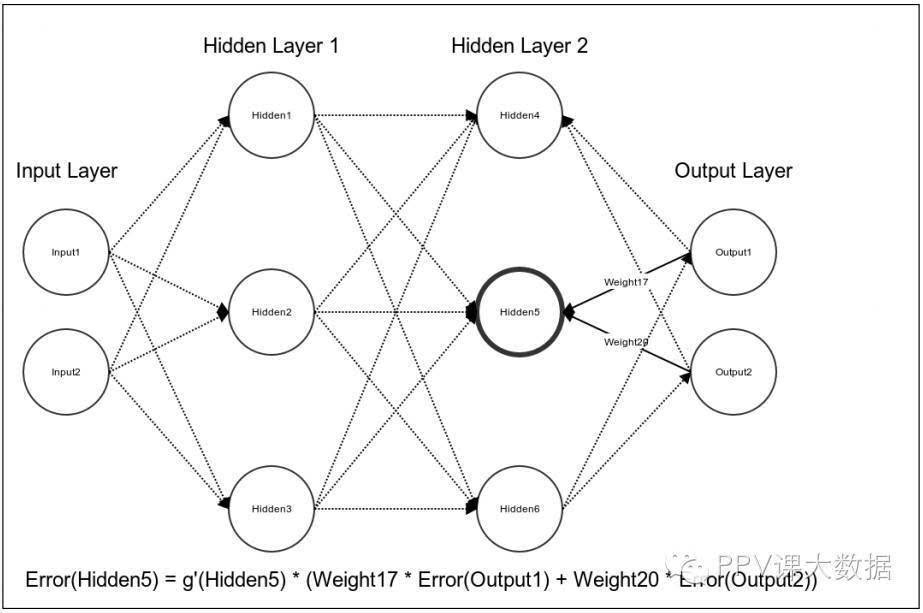
As a very effective mathematical work, machine learning algorithms, especially neural networks, have appeared as early as the last practical 1960s. At that time, scientists took biological neural networks as the prototype and put forward the first generation of neural network model, which was called perceptron at that time.[4] The characteristics of the first generation of neural network are that its characteristics need to be input manually in advance. In the process of pattern recognition, It is based on feature weight. Its limitation lies in the poor effect of dealing with specific tasks. Therefore, the application of the first generation neural network is often very limited. The first generation neural network perceptron model is as follows.

##### 1

##### Figure 1-The first generation neural network perceptron model (from Google)

Marvin Minsky and Seymour papert discussed the shortcomings of the first generation neural network in 1969 [5], which made other researchers less and less confident about the neural network. Therefore, since 1970, the research and development of neural network has entered a low tide.

In 1982, physicist John Hopfield proposed a new neural network model, called Hopfield model, and proposed the definition of energy for the first time. Hopfile model has strong fault-tolerant performance, and can reconstruct complete data image from incomplete or distorted data image. After the first generation of neural network model, Hopfield's model is called the second generation of neural network. Different from the first generation of perceptron, Hopfield's model no longer uses the preset features, but is automatically adapted by the algorithm in the training process, so it solves the problem of narrow range of use. For training suitability, this kind of neural network adopts a back propagation algorithm [6][7], which propagates the error of the sample back to the input layer from the input layer. The weight of the network is constantly adjusted in the process of propagation to improve the likelihood of the neural network for the sample. At the same time, although the hidden layer can enhance the expression ability of the network, its introduction also makes the network more complex. The second generation neural network model is shown in the figure.



##### Figure 2-The second generation neural network model(from Google)

Compared with the first generation neural network, the second generation can deal with many problems, but it has obvious disadvantages: [7]

1. supervised learning is its only purpose, and it needs to use label when training data. The vast majority of real-world data is unmarked.
2. The training time is very long, especially when there are many hidden layers, but if the hidden layer is reduced, the accuracy will be reduced.
3. Back propagation algorithm is easy to fall into local optimal solution.

Therefore, the development of neural network has entered the second trough.

Until 2016, Hinton first proposed the theory of deep network and deep learning, and published it in science, which attracted people's attention to deep learning. Deep learning architecture is composed of multi-layer nonlinear operation units, which can learn complex functions representing high-order abstract concepts. In the process from low to high, the bottom output is used as the high-level input. Through many times of input data learning, we can get the high-order feature representation of its structure information, which can be used in specific problems such as classification, regression and information retrieval. The birth of deep learning gives new hope to the study of artificial intelligence [7].

However, at the present research level, people still can't grasp the characteristics of each layer. In this design, GSP method is proposed to analyze the characteristics of the middle layer of the deep neural network, so as to provide the basis for the accuracy and performance of the deep neural network.

## GSP Application:

The application of network or graph is very common in scientific research and daily life, such as social network, biological network, communication network and so on. In recent years, with the development of artificial intelligence, deep neural network is also rising. However, deep learning has always been presented in the form of "black box" in front of researchers. Although many deep neural network structures have made unprecedented achievements in various fields recently, such as the field of image classification, the accuracy of deep neural network can reach more than 70% and more than 90%, which has reached or even exceeded the human score Class ability. However, there is no clear theory to support why the deep neural network can achieve such accuracy. In fact, most of the neural network construction or innovation often have guesswork components, so it is urgent to find a way to go deep into the neural network to find some indicators that can measure or guide the neural network.

The essence of deep neural network is directed acyclic graph. Each node represents a data node that can be associated with a label, and the graph is formed by connecting different node edges and assigning each edge its own weight. Since it is a graph, it must follow graph theory, so the methods related to graph theory have very important guiding significance in the application of deep neural network. Graph signal processing (GSP) is one of the most important tools in graph theory. GSP can process and filter different graph nodes in different forms. When data labels are presented as signals on the graph, the graph signal regularization technique can be used in the process of label estimation to optimize the prediction of unknown labels in classification and semi supervised learning problems.

# Lit Review:

## Report on GSP for Semi-Superviesd Learning:[x1]

In 2014 a report was published which introduced a novel framework for bacth mode active semi-supervised learning based on sampling theory for graph signals. The proposed active learning framework aims to select the subset nodes which maximizes the dimension of the space of uniquely recoverable signals. In the context of sampling theory, this translates to selecting the subset with the maximum cut-oﬀ frequency. This interpretation leads to a very eﬃcient greedy algorithm. They provide intuition about how the method tries to choose the nodes which are most representative of the data. they also present an eﬃcient semi-supervised learning method based on bandlimited interpolation. they show, through experiments on real data, that our two algorithms, in conjunction, perform very well compared to state of the art methods. they used their proposed activate semi-supervised learning algorithm to perform a calssification task on several popular dataset to cimpare with other state of the art methods. The results are ilustrated in Figure 3.

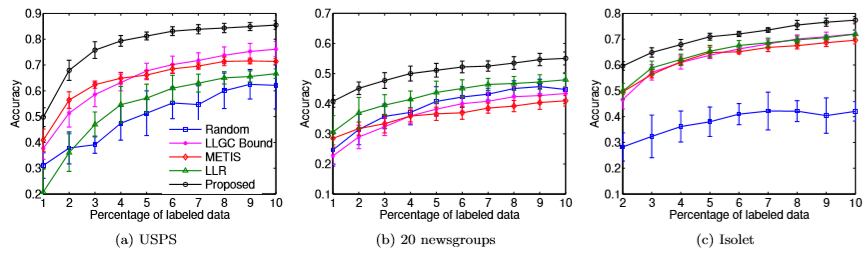


Figure 3 [x1]. Comparison of active semi-supervised learning methods on real datasets. Plots show the average classiﬁcation accuracies for diﬀerent percentages of labeled data

They repeat the classiﬁcation over 10 such instances of the dataset and report the average classiﬁcation error. The results are illustrated in Figure (3a). they observe that their proposed method outperforms the others. A notable feature of their method is that they show very good classiﬁcation results even for very few labeled samples. This is due to their inherent criterion for active learning that tries to select those points that maximize the recoverable dimensions of the underlying data manifold. The classiﬁcation results in Figure (3b) show that their method performs very well compared to others. However, the absolute error rates are not very good. This is due to the high similarity between diﬀerent newsgroups which makes

the problem inherently diﬃcult. The experiment is repeated over 10 instances of the dataset and average prediction error is reported in Figure (3c). Note that they start with 2% labeled points to ensure that each method gets a fair chance of selecting at least one point to label from each of the 26 classes. they observe that their method outperforms the others.

## Report on Fourier Analysis of DNNs:[x2]

Another report on the Fourier analysis of Deep Neural Network wa published in 2018. They studied deep ReLU networks through the lens of Fourier analysis. Several conclusions can be drawn from their analysis. While neural networks can approximate arbitrary functions, they ﬁnd that they favour low frequency ones – hence they exhibit a bias towards smooth functions – a phenomenon that was called spectral bias. they also illustrated how the geometry of the data manifold impacts expressivity in a non-trivial way, as high frequency functions deﬁned on complex manifolds can be expressed by lower frequency network functions deﬁned in input space. From the figure 4, they ﬁnd that even when higher frequencies have larger amplitudes, the model prioritizes learning lower frequencies ﬁrst. they also ﬁnd that the spectral norm of weights increases as the model ﬁts higher frequency, which is what they expect .

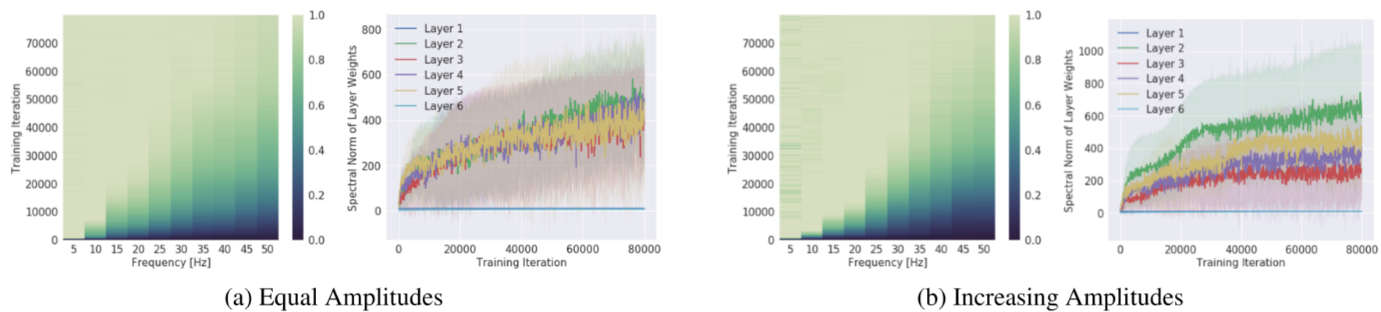


figure 4 [x2]. Amplitude and frequency comparison

Left (a, b): Evolution of the spectrum (x-axis for frequency) during training (y-axis). The colors show the measured amplitude of the network spectrum at the corresponding frequency, normalized by the target amplitude at the same frequency and the colorbar is clipped between 0 and 1. Right (a, b): Evolution of the spectral norm (y-axis) of each layer during training (x-axis). Figure-set (a) shows the setting where all frequency components in the target function have the same amplitude, and (b) where higher frequencies have larger amplitudes.

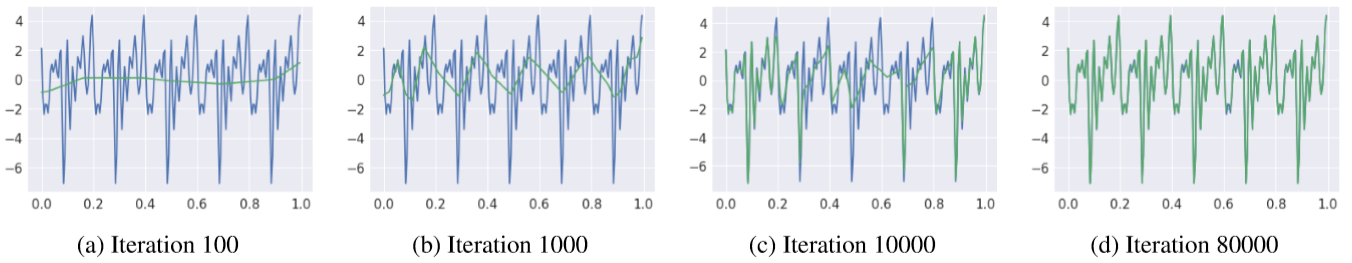


figure5 [x2]The learnt function (green) overlayed on the target function (blue) as the training progresses. The target function is a superposition of sinusoids of frequencies κ = (5,10,...,45,50), equal amplitudes and randomly sampled phases.

Figure 5 shows the learned function at intermediate training iterations. The result is that lower frequencies are regressed ﬁrst, regardless of their amplitudes.

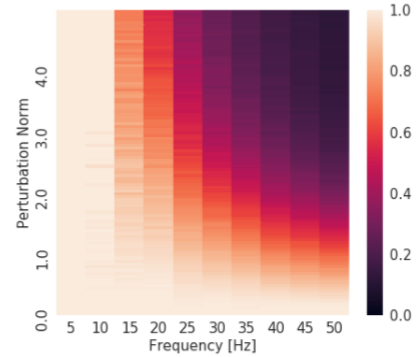


figure 6 [x2]. Normalized spectrum of the model (x-axis for frequency, colorbar for magnitude)with perturbed parameters as a function of parameter perturbation (y-axis). The colormap is clipped between 0 and 1. We observe that the lower frequencies are more robust to parameter perturbations than the higher frequencies.

The result, shown in Figure 6, demonstrates that higher frequencies are signiﬁcantly less robust than the lower ones, guiding the intuition that expressing higher frequencies requires the parameters to be ﬁnely-tuned to work together. In other words, parameters that contribute towards expressing high-frequency components occupy a small volume in the parameter space

## Report on Predicting DNNs Over fitting using GSP:[x3]

another report on predicting DNNs overfitting using GSP was published in 2018. they have shown via experiments that there exists a strong correlation between smoothness gap and generalization abilities in deep neural networks.

Deep Neural Networks (DNNs) have become the state-ofthe-art in many machine learning benchmarks ever since the AlexNet won the ILSVRC-2012 competition. Due to the fact they rely on millions of trainable parameters, DNNs remain black-box methods. As a consequence, there is little understanding of the reasons for their generalization abilities and finding the best hyperparameters for a given problem requires to exhaustively search a lot of combinations. It is often considered that architectures that are not optimal for a given problem suffer either from what is called underfitting or overfitting. Underfitting typically refers to conditions where increasing the number of trainable parameters in some parts of the architecture would lead to better generalization performance. To the contrary, overfitting refers to conditions where the network is containing too many parameters and despite being very efficient at classifying the training set, it fails when facing unseen examples.

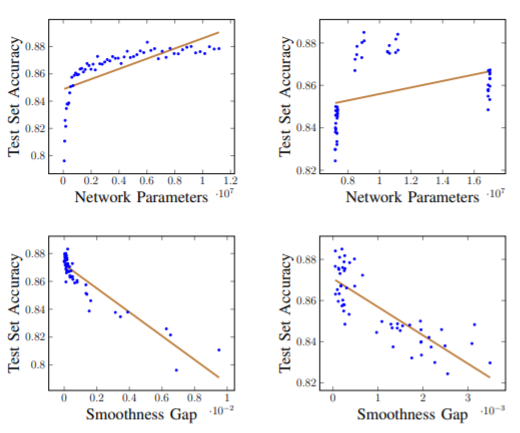


figure 7 [x3]. Results generated by varying the size of the dataset

Results are depicted in figure 7, The right column shows results where the size of feature maps is changed independently, while the left column the feature maps depend on the size of the first convolutional layer. While the upper row shows the correlation between the number of trainable parameters and the network performance, while the lower row shows the correlation between the smoothness gap and network performance. The brown lines are linear regressions.

# Work Plan and Gantt CHART:

The overall work plan for this project can be broken down into the following objectives:

1. Choose a suitable Deep Neural Network and corresponding data set.
2. Training the Deep Neural Network to the excepted accuracy.
3. Compare and choose several feature of the intermediate representations of above Deep Neural Network.
4. Characterize the above intermediate representations using graph signal processing(GSP).
5. Analysis the performance of the above experiments and research the correlation between over-fitting and DNN architecture and so on.

In the first step, we first need to learn and summarize the knowledge related to deep neural network, deep learning and graph signal processing, which is the basis of completing our whole project. Deep learning is mainly to simulate the learning process of human brain through artificial neural network (ANN), which is a promising branch of machine learning. Image classification technology is a kind of technology to process image information by computer, but the result of image processing is quite different between human and computer. Artificial intelligence methods such as deep learning can analyze and process image data more rigorously and scientifically.

In recent years, how to quickly and effectively extract and analyze the semantic information contained in these images and apply it to practical problems has become the research focus of image classification and recognition, image search, image understanding and analysis and other fields. With the development of image research, intelligent image classification algorithms such as fuzzy set method, decision tree classification method, knowledge-based classification method and machine learning method are emerging. As an important branch of machine learning, deep learning algorithm [1] provides a new direction for the research of image classification technology. The deep learning algorithm simulates the multi-layer neural network to establish the model according to the characteristics that the human visual system processes the information hierarchically.

The original intention of deep learning research is to apply it to image recognition. So far, although deep learning has been applied to speech, image, text, emotion and other fields, most of the papers published in the field of deep learning are about image recognition. Deep learning has great advantages in image recognition and classification. At present, deep learning algorithm has become the mainstream of machine learning and has a broad research value and application prospect in many fields, such as image classification, speech recognition and so on. In-depth learning, unsupervised learning is used, and the sample label is unknown in the training process. This learning method does not need human participation. In the modern information age, large sample and large data sets are becoming more and more common. Deep learning can use multi-layer nonlinear transformation to process a large number of unlabeled data such as images, sounds, and texts, so as to achieve supervised or unsupervised feature extraction and conversion, pattern analysis and classification.

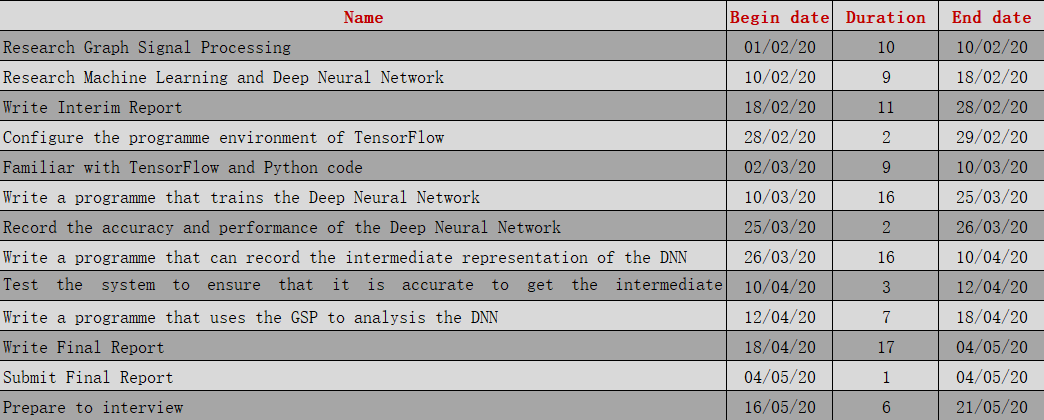
The second step is to train a suitable deep neural network structure. At present, researchers have developed a large number of targeted and complex deep neural network models to solve different problems, the most typical of which is convolutional neural network, so this project will also be based on convolutional neural network for research.

The third step is to choose a suitable data set. At present, there are a large number of different data sets in the whole field that can be used for free, among which CFAIR-10 data set is the most classic. CFAIR-10 data set is a data set used for image classification, and its image types are divided into 10 kinds. Each kind of image has more than 5000 pictures, a total of more than 50000 pictures, which is enough for the training of general depth neural network. In order to verify the training effect of the deep neural network, 2 / 3 of the whole data set is generally regarded as the training set, and the rest 1 / 3 as the test set.

The fourth step is to select the appropriate features. At present, there are some mainstream features proposed for the graph signal processing in the scientific research field. Among them, K-Nearest Neighbor Graphs will be the main feature of my research. In the later stage, in the actual operation process, some new features may be added.

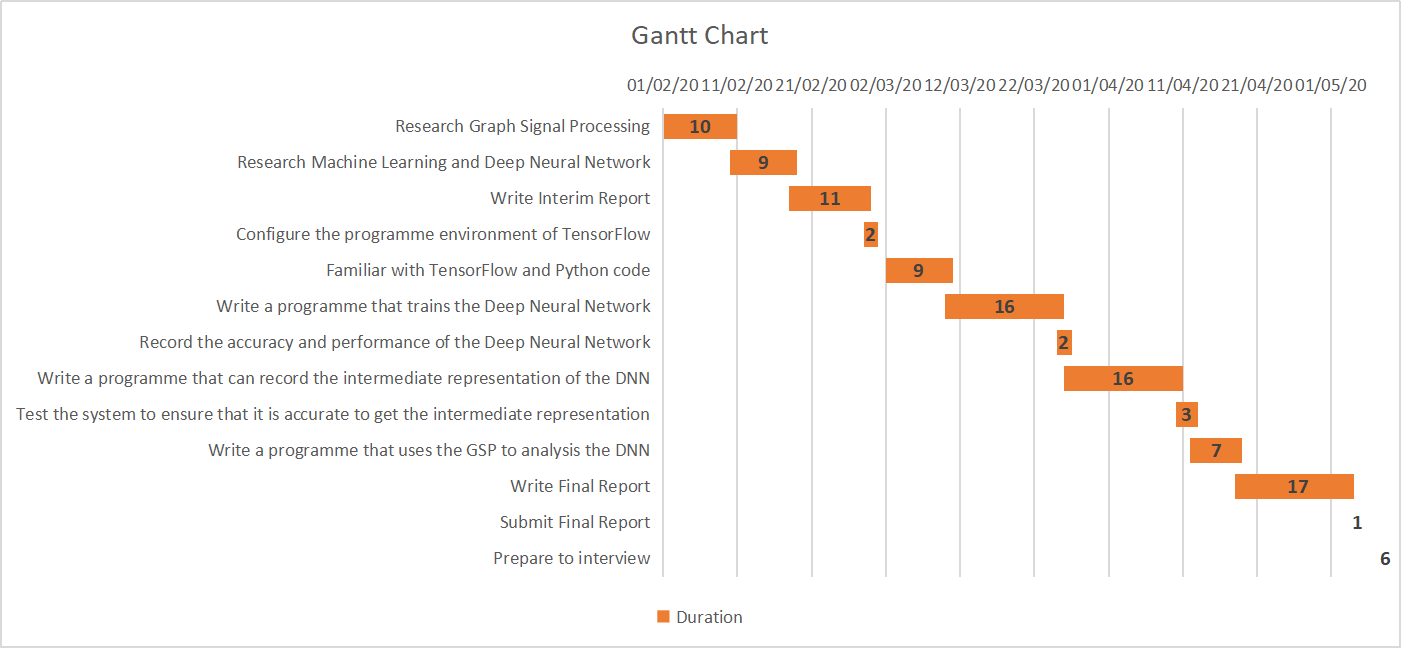
The fifth and most complex step is to analyze the relationship between the different performance of deep neural network and its characteristics, so as to provide possible guidance for the training of deep neural network.

The programming related to the whole project will be developed based on Tensorflow, the most popular deep learning framework at present. In addition, the programming related to GSP is supported by PyGSP module in Python.



##### Figure -Work plan

##### Figure -Gantt chart



# References

1. S. Mallat, “Understanding deep convolutional networks,” Phil. Trans. R. Soc. A, vol. 374, no. 2065, p. 20150203, 2016.
2. J. Viswanathan, F. Rémy, N. Bacon-Macé, and S. J. Thorpe, “Long term memory for noise: evidence of robust encoding of very short temporal acoustic patterns,” Frontiers in neuroscience, vol. 10, p. 490, 2016.
3. C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, “Intriguing properties of neural networks,” arXiv preprint arXiv:1312.6199, 2013.
4. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain [J]. Psychological review, 1958, 65(6): 386.
5. Minsky M, Papert S. Perceptron: an introduction to computational geometry [J]. The MIT Press, Cambridge, expanded edition, 1969, 19(88): 2.
6. Hopfield J J.Neural networks and physical systems with emergent collective computational abilities [J]. Proceedings of the national academy of sciences, 1982, 79(8): 2554-2558.
7. Arel I, Rose D C, Karnowski T P. Deep machine learning-a new frontier in artificial intelligence research [research frontier][J]. Computational Intelligence Magazine, IEEE, 2010, 5(4): 13-18.

[x1] A. Gadde, A. Anis, and A. Ortega, “Active semi-supervised learning using sampling theory for graph signals,” in Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2014, pp. 492–501.

[x2] Rahaman N , Baratin A , Arpit D , et al. On the Spectral Bias of Neural Networks[J]. 2018.

[x3] Lassance C E R K, Gripon V, Ortega A. Predicting Under and Overfitting in Deep Neural Networks using Graph Smoothness[J].